V9: Cellular differentiation - Epigenetics

In developmental biology, **cellular differentiation** is the process where a cell changes from one cell type to another.
Most commonly the cell changes to a **more specialized type**.

Differentiation occurs numerous times during the development of a multicellular organism as it changes from a simple zygote to a complex system of tissues and cell types.

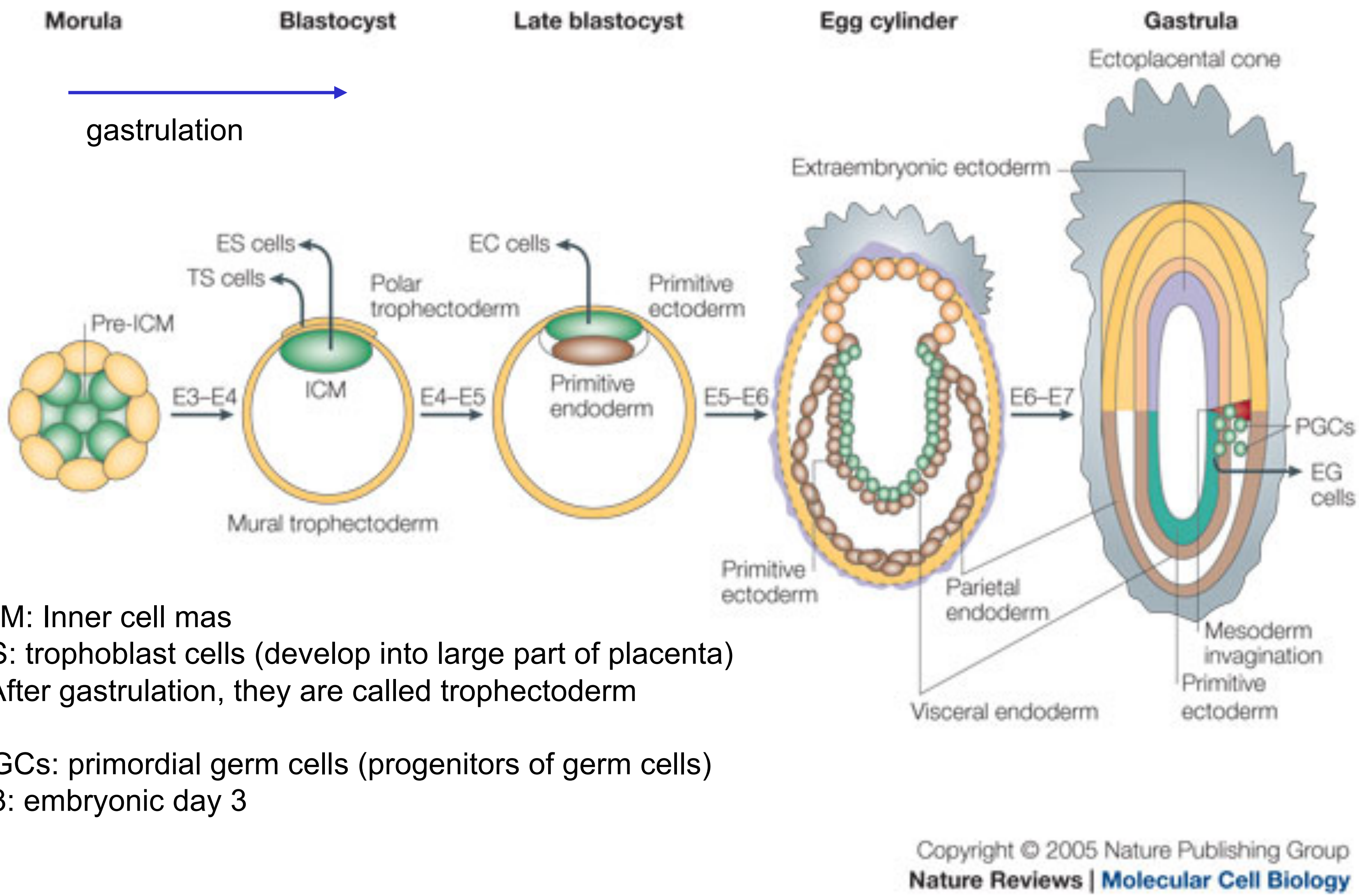
Differentiation continues in **adulthood** as adult stem cells divide and create fully differentiated daughter cells during tissue repair and during normal cell turnover.

Differentiation dramatically changes a cell's size, shape, membrane potential, metabolic activity, and responsiveness to signals.

These changes are largely due to highly controlled modifications in gene expression that are often controlled by **epigenetic** effects.

www.wikipedia.org

Embryonic development of mouse



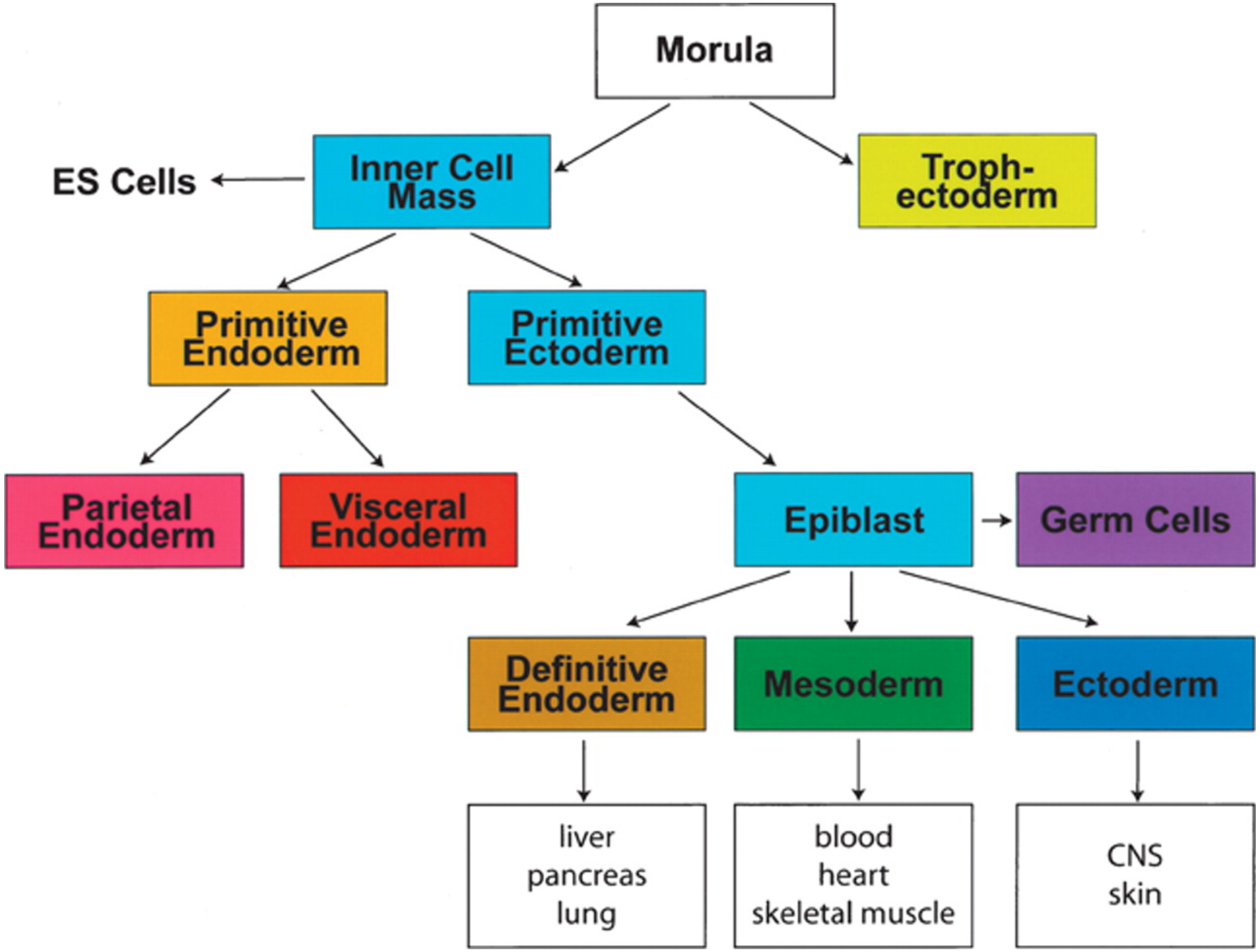
gastrulation

ICM: Inner cell mas
TS: trophoblast cells (develop into large part of placenta)
- After gastrulation, they are called trophectoderm

PGCs: primordial germ cells (progenitors of germ cells)
E3: embryonic day 3

Boiani & Schöler, Nat Rev Mol Cell Biol 6, 872 (2005)

Cell populations in early mouse development



Scheme of early mouse development depicting the relationship of early cell populations to the primary germ layers

Keller, Genes & Dev. (2005) 19: 1129-1155

Types of body cells

3 basic categories of cells make up the mammalian body:

germ cells (oocytes and sperm cells)
somatic cells, and
stem cells.

Each of the approximately 100 trillion (10^{14}) cells in an adult human has its own copy or copies of the genome except certain cell types, such as red blood cells, that lack nuclei in their fully differentiated state.

Most cells are **diploid**; they have two copies of each chromosome.

Cells differentiate to specialize for different functions.

Somatic cells make up most of the human body, such as skin and muscle cells.

www.wikipedia.org

Epigenetic landscape during early development

Embryonic development is a complex process that remains to be understood despite knowledge of the complete genome sequences of many species and rapid advances in genomic technologies.

A fundamental question is how the unique gene expression pattern in each cell type is established and maintained during embryogenesis.

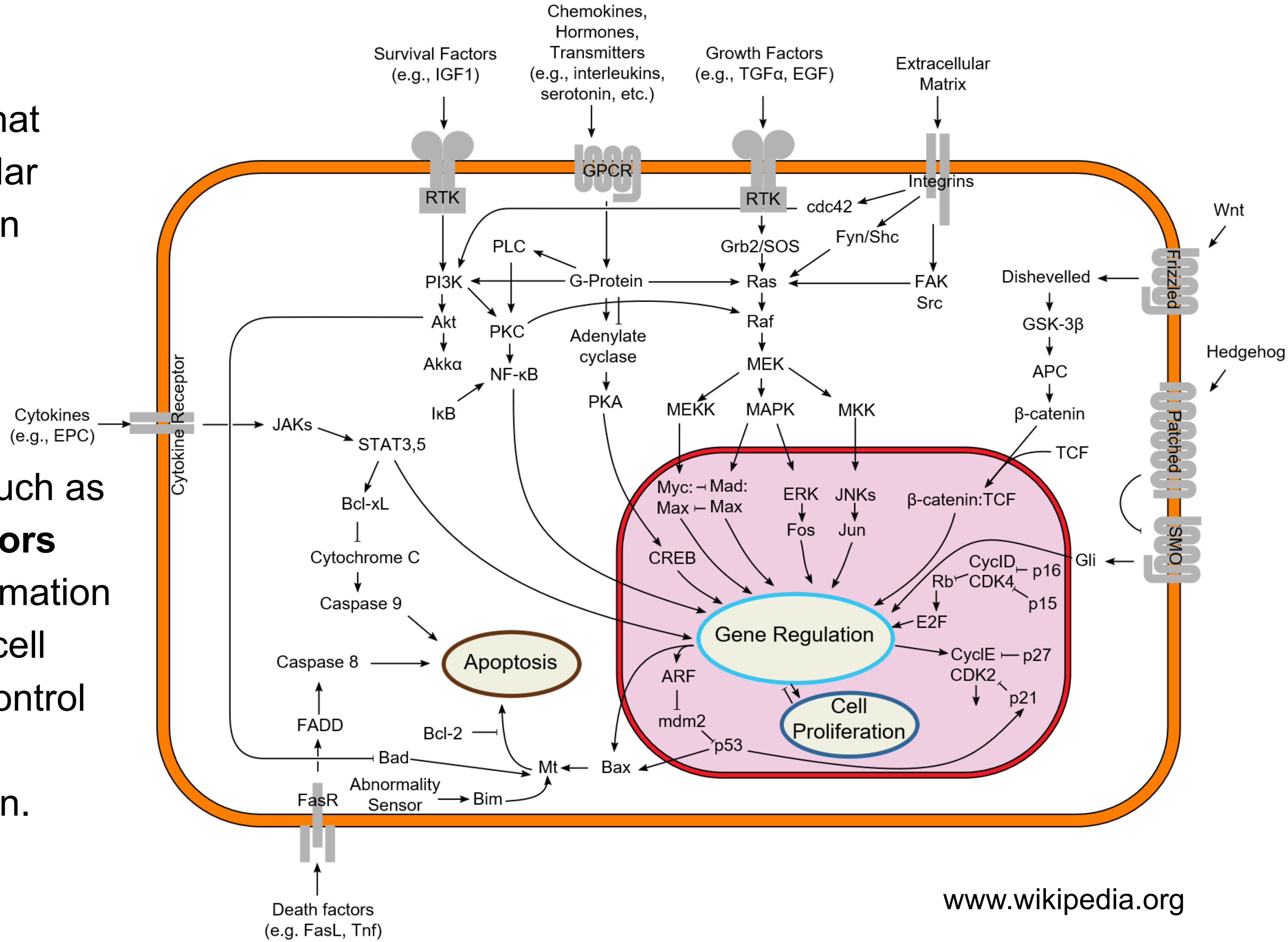
It is well accepted that the gene expression program encoded in the genome is executed by **transcription factors** that bind to cis-regulatory sequences and modulate gene expression in response to **environmental cues**.

Xie et al., Cell 153, 1134-1148 (2013)

Growth factors induce cell differentiation

The major molecular processes that control cellular differentiation involve **cell signaling**.

Signalling molecules such as **growth factors** convey information from cell to cell during the control of cellular differentiation.



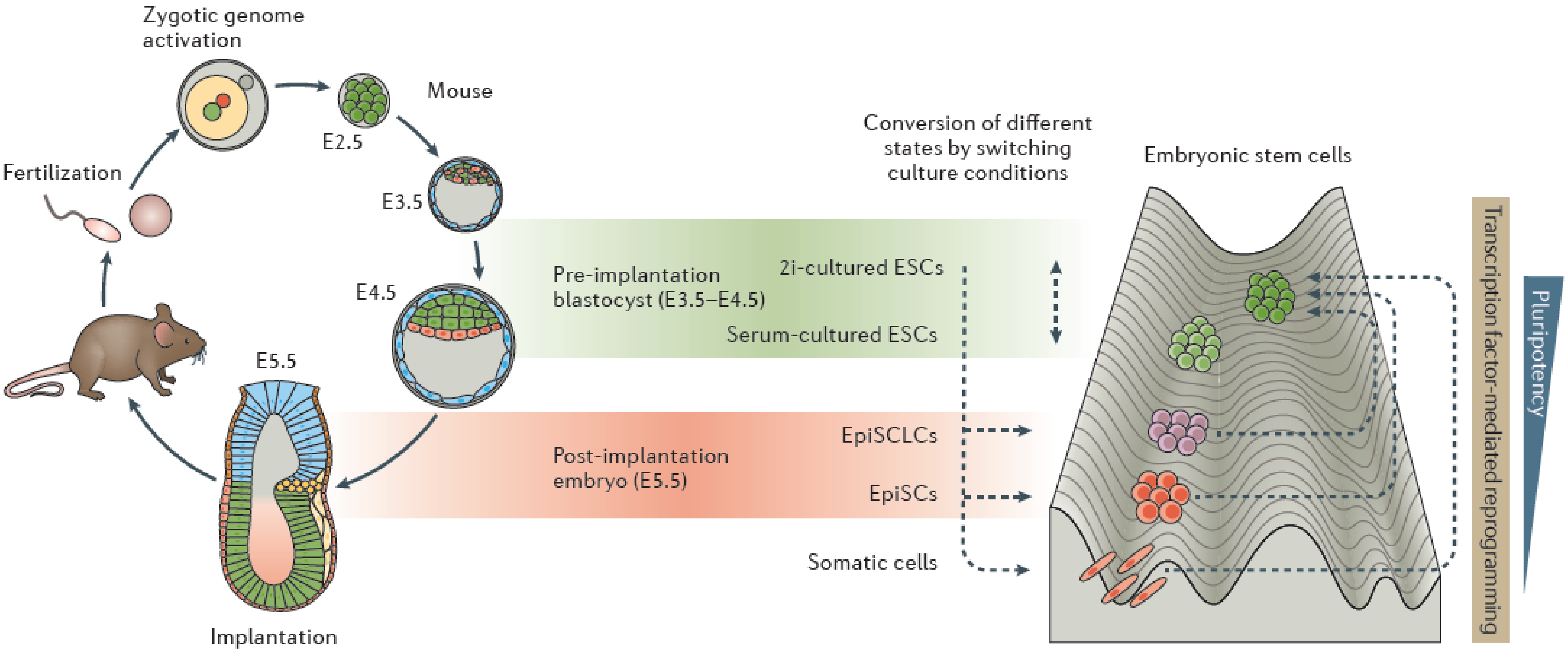
www.wikipedia.org

Different states of pluripotency

E4.5 epiblast cells: represent ground-state pluripotency

Implantation: stage of pregnancy at which the blastocyst adheres to the wall of the **uterus**.

After implantation (E5.5): **epiblast cells** undergo a strong wave of epigenetic reprogramming. They are now „primed“.



Atlasi & Stunnenberg, *Nature Rev Genet* **18**, 643–658 (2017)

DNA methylation data

Epigenetics refers to **alternate phenotypic states** that are **not based** on **differences in genotype**, and are potentially reversible, but are generally stably maintained during cell division.

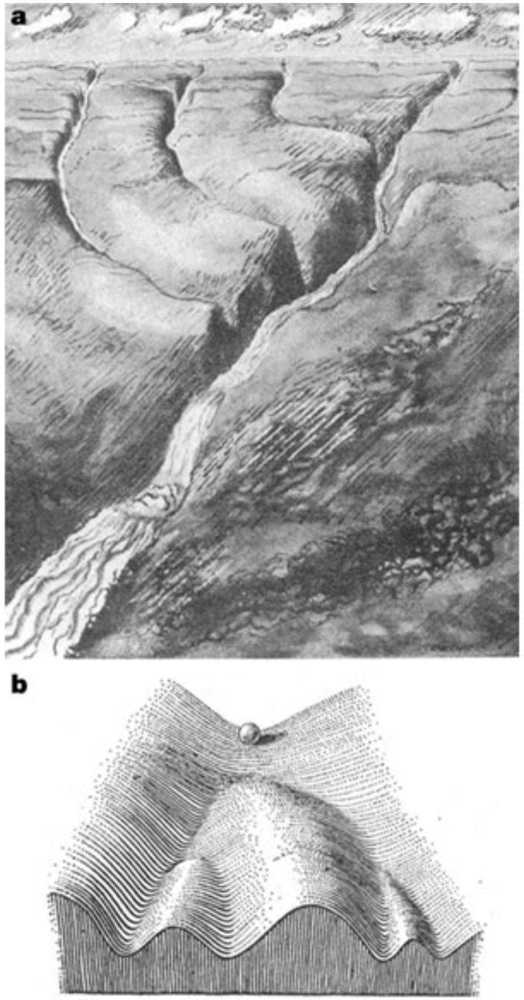
<u>Examples</u>: imprinting, twins, cancer vs. normal cells, differentiation, ...

Multiple mechanisms interact to collectively establish

- alternate states of chromatin structure (open – packed/condensed),
- **histone modifications**,
- composition of associated proteins (e.g. histones),
- transcriptional activity,
- activity of microRNAs, and
- in mammals, **cytosine-5 DNA methylation** at CpG dinucleotides.

Laird, Hum Mol Gen 14, R65 (2005)

Waddington's epigenetic landscape for embryology



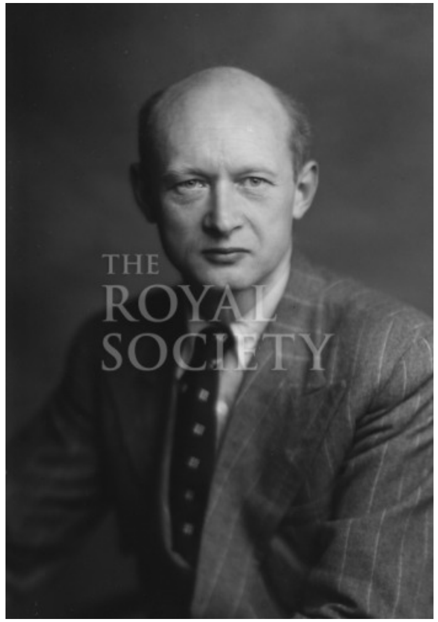
Slack, Nature Rev Genet 3, 889-895 (2002)

Waddington worked in **embryology**

a) is a painting by John Piper that was used as the frontispiece for Waddington's book *Organisers and Genes*.
It represents an epigenetic landscape.

Developmental pathways that could be taken by each cell of the embryo are metaphorically represented by the path taken by water as it flows down the valleys.

b) Later depiction of the epigenetic landscape. The ball represents a cell, and the bifurcating system of valleys represents bundles of trajectories in state space.



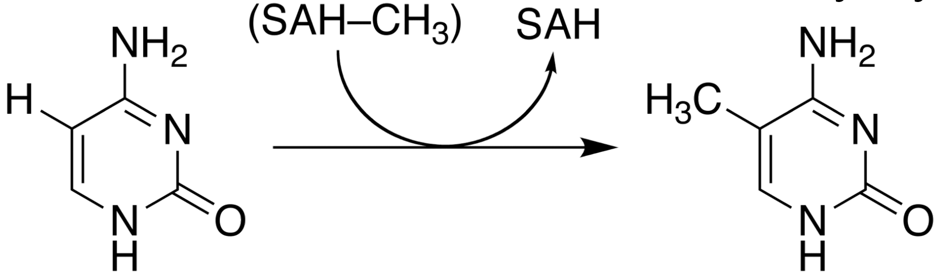
Conrad Hal Waddington (1905 – 1975)
pictures.royalsociety.org

Cytosine methylation

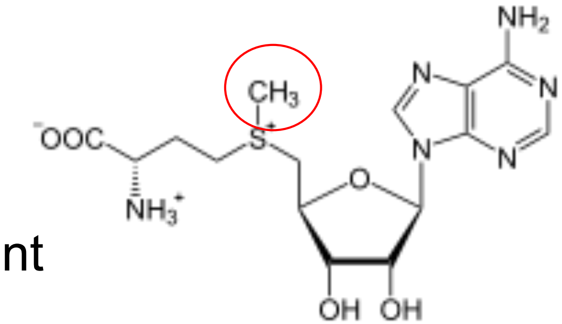
<u>Observation</u>: 3-6 % of all cytosines are methylated in human DNA.
This methylation occurs (almost) exclusively when cytosine is followed by a guanine base -> **CpG dinucleotide**.

Cytosine

5-methyl-cytosine



SAM: S-adenosyl-methionine
SAH: S-adenosyl-homocysteine



Mammalian genomes contain much fewer (only 20-25 %) of the CpG dinucleotide than is expected by the G+C content (we expect 1/16 ≈ 6% for any random dinucleotide).

This is typically explained in the following way:

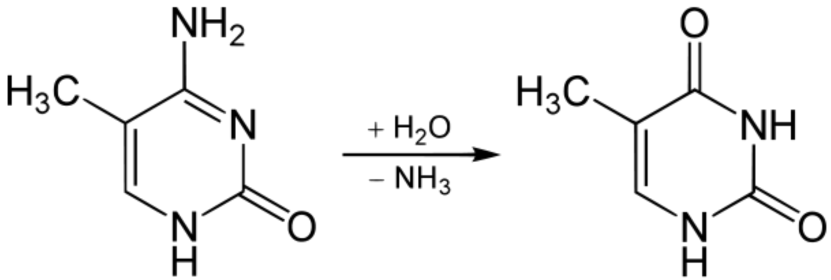
As most CpGs serve as targets of DNA methyltransferases, they are usually methylated …. (see following page)

Esteller, Nat. Rev. Gen. 8, 286 (2007)
www.wikipedia.org

Cytosine methylation

But 5-Methylcytosine can easily **deaminate** to **thymine**.

5-methyl-cytosine



thymine

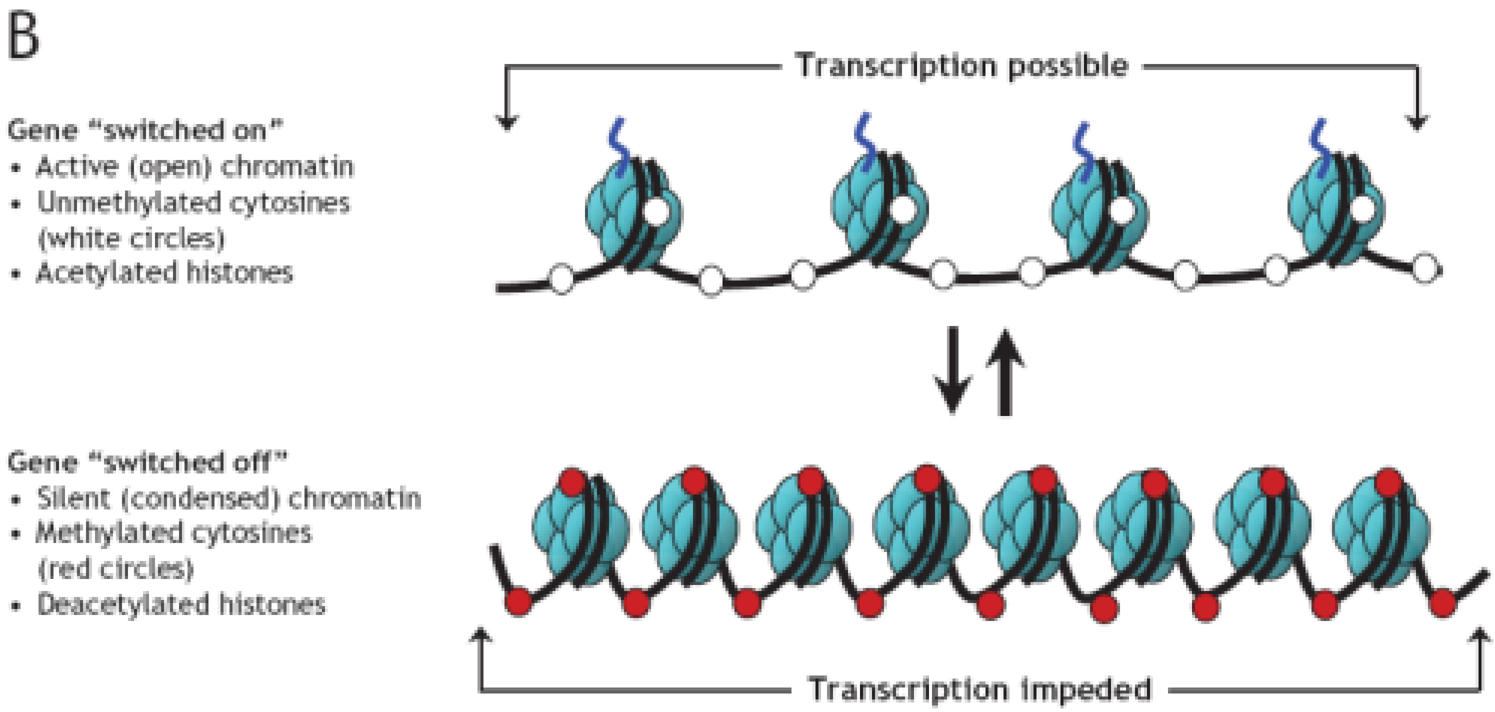
If this mutation is not repaired, the affected CpG is permanently converted to TpG (or CpA if the transition occurs on the reverse DNA strand).

Hence, methylCpGs represent **mutational hot spots** in the genome.
If such mutations occur in the germ line, they become heritable.

A constant loss of CpGs over thousands of generations
can explain the low frequency of this
special dinucleotide in the genomes of human and mouse.

Esteller, Nat. Rev. Gen. 8, 286 (2007)
www.wikipedia.org

chromatin organization affects gene expression



Schematic of the reversible changes in chromatin organization that influence gene expression:
genes are expressed (switched on) when the chromatin is **open** (active), and they are inactivated (switched off) when the chromatin is **condensed** (silent).

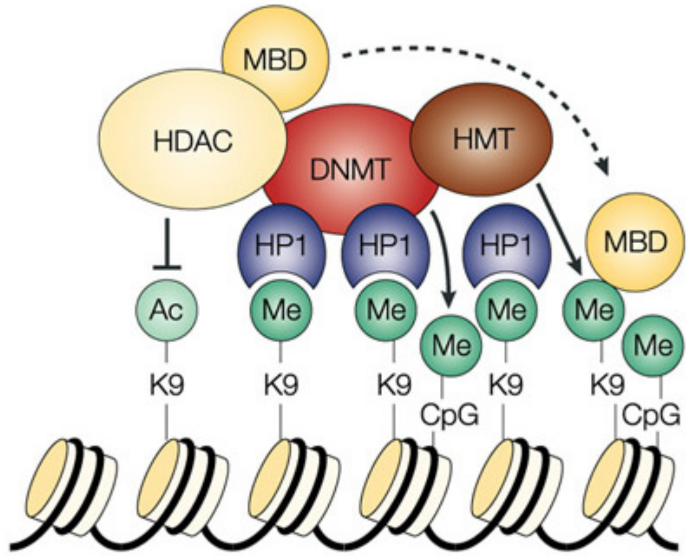
White circles = unmethylated cytosines;
red circles = methylated cytosines.

Rodenhiser, Mann, CMAJ 174, 341 (2006)

Enzymes that control
DNA methylation and histone modfications

These dynamic chromatin states are controlled by reversible epigenetic patterns of **DNA methylation** and **histone modifications**.

Enzymes involved in this process include
- DNA methyltransferases (DNMTs),
- histone deacetylases (HDACs),
- histone acetylases,
- histone methyltransferases (HMT) and the
- methyl-binding domain protein MECP2 with its methyl-binding domain (MBD) that binds specifically to me-cytosine.



HP1: heterochromatin protein 1

Rodenhiser, Mann, CMAJ 174, 341 (2006)
Feinberg AP & Tycko P (2004) Nature Reviews: 143-153

DNA methylation

Typically, unmethylated clusters of CpG pairs are located in **tissue-specific genes** and in essential **housekeeping genes.**

(House-keeping genes are involved in routine maintenance roles and are expressed in most tissues.)

These clusters, or **CpG islands**, are targets for proteins that bind to unmethylated CpGs and initiate gene transcription.

In contrast, **methylated CpGs** are generally associated with silent DNA, can block methylation-sensitive proteins and can be easily mutated.

The **loss** of normal DNA methylation patterns is the best understood epigenetic cause of **disease**.

In animal experiments, the removal of genes that encode DNMTs is lethal; in humans, overexpression of these enzymes has been linked to a variety of cancers.

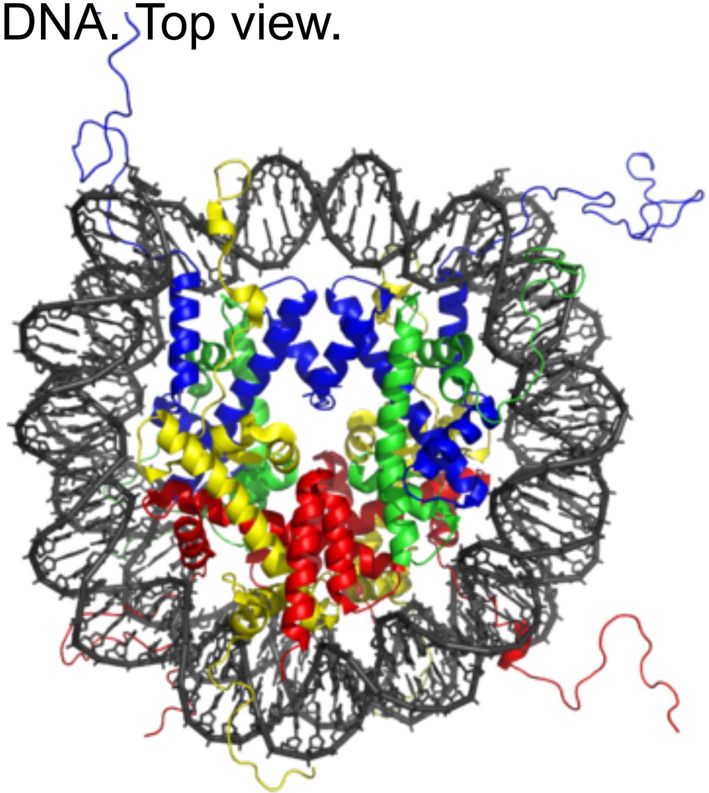
Rodenhiser, Mann, CMAJ 174, 341 (2006)

The histone code

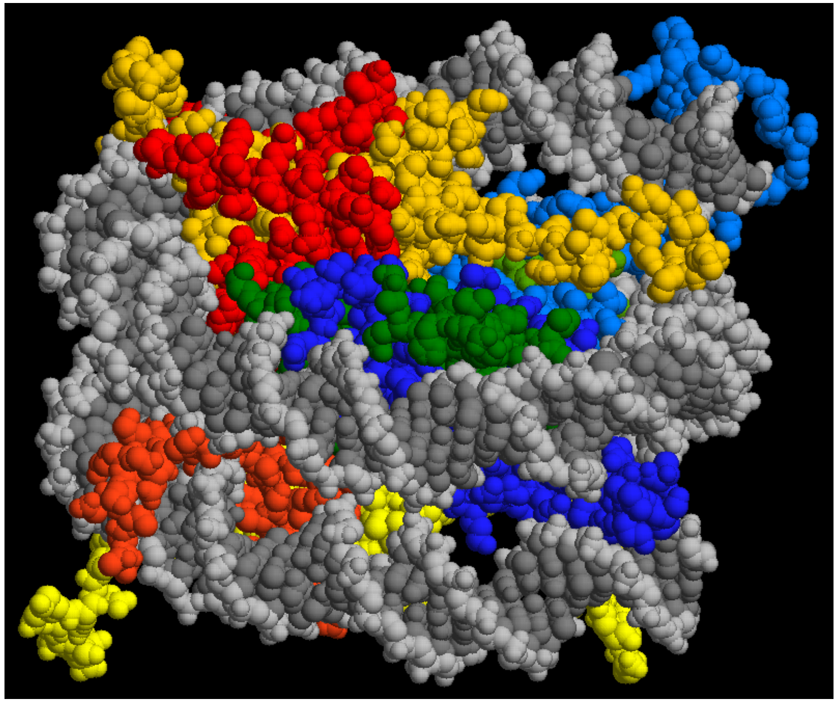
The DNA of eukaryotic organisms is packaged into chromatin, whose basic repeating unit is the **nucleosome**.

A nucleosome is formed by wrapping 147 base pairs of DNA twice around an octamer of four core histones, **H2A** , **H2B** , **H3** and **H4** (2 copies of each one).

X-ray structure of the nucleosome core particle consisting of core histones, and DNA. Top view.



Side view shows two windings of DNA and two histone layers



www.wikipedia.org

Post-translational modifications of histone tails

*ACETYLATION AND METHYLATION OF HISTONES AND THEIR POSSIBLE ROLE IN THE REGULATION OF RNA SYNTHESIS**

BY V. G. ALLFREY, R. FAULKNER, AND A. E. MIRSKY

THE ROCKEFELLER INSTITUTE

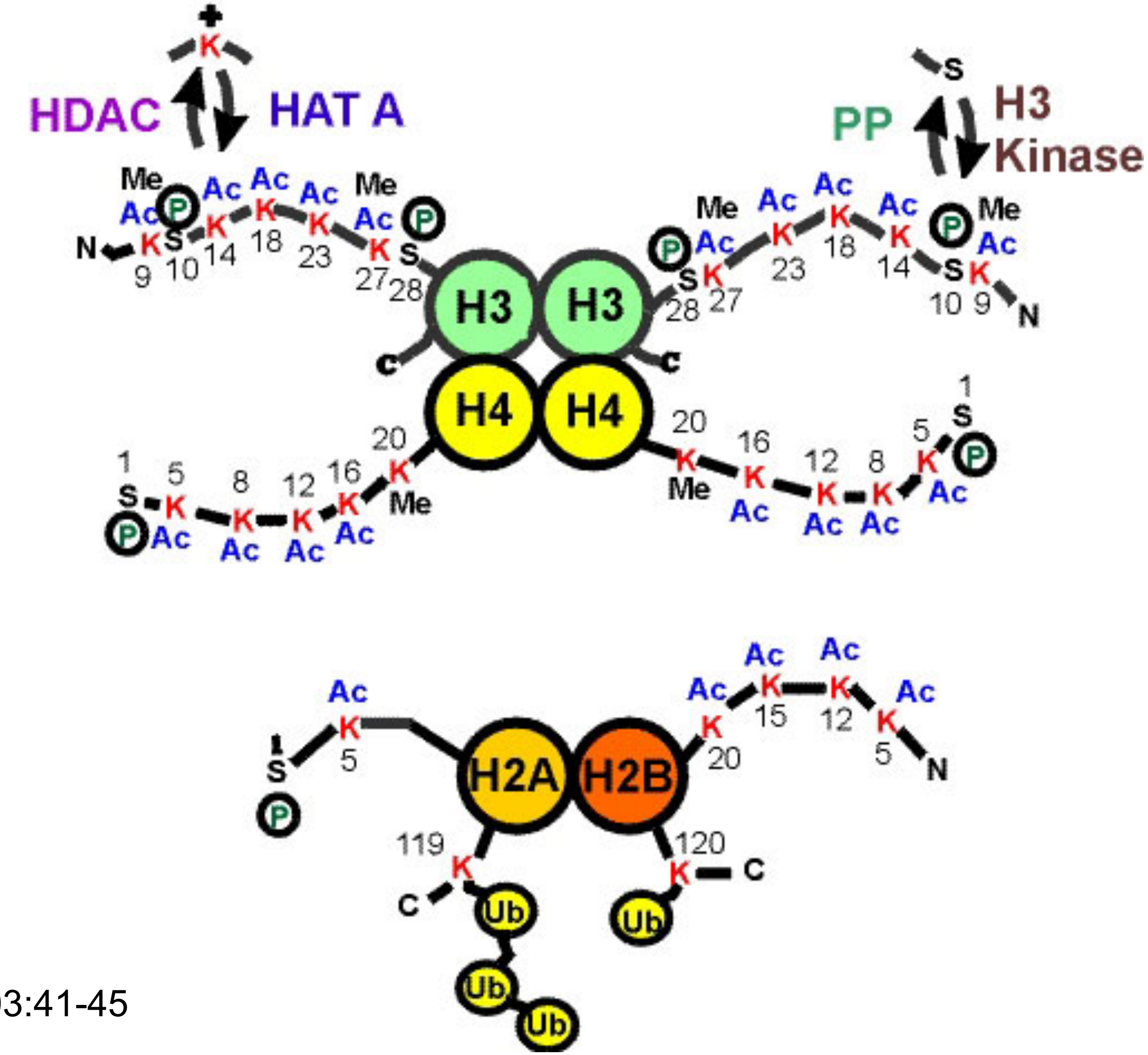
PNAS 1964;51:786
First report on PTMs of histones

The disordered histone tails comprise 25-30% of the histone mass.

They extend from the compact histone multimer to provide a platform for various **post-translational modifications (PTMs)**.

These modifications affect the histones' ability to bind DNA and to other histones.

This, in turn, affects **gene expression.**



Strahl BD and Allis CD, 2000. Nature 403:41-45

Mode of action of histone PTMs

Histone PTMs exert their effects via two main mechanisms.

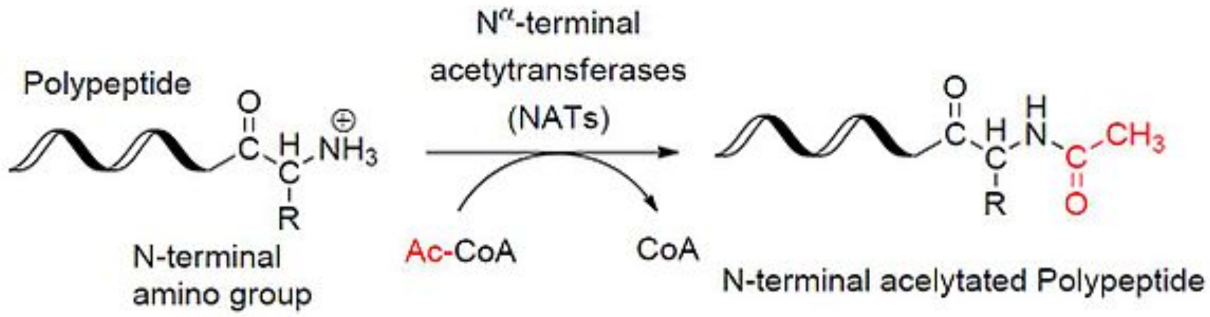
(1) PTMs directly influence the overall structure of chromatin, either over short or long distances.

(2) PTMs regulate (either positively or negatively) the binding of effector molecules.

Bannister, Kouzarides, Cell Res. (2011) 21: 381–395.

PTMs of histone tails

Histone **acetylation** and **phosphorylation** effectively reduce the positive charge of histones.



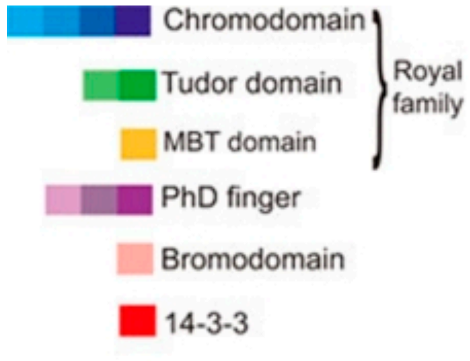
This potentially disrupts electrostatic interactions between histones and DNA.

This presumably leads to a less compact chromatin structure, thereby facilitating DNA access by protein machineries such as those involved in transcription.

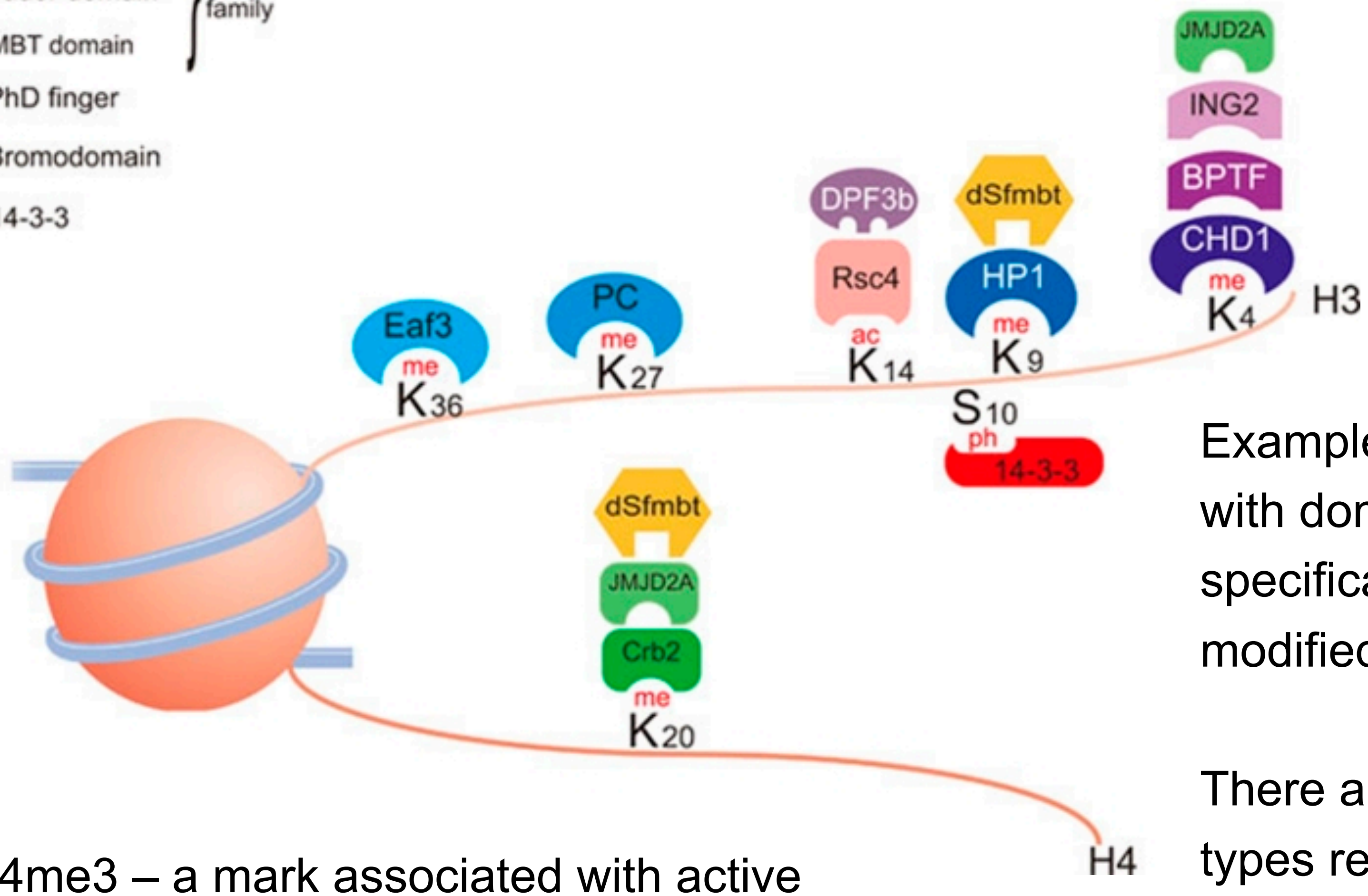
Histone **methylation** mainly occurs on the side chains of lysines and arginines.

Unlike acetylation and phosphorylation, however, histone methylation does not alter the charge of the histone protein.

Bannister, Kouzarides, Cell Res. (2011) 21: 381–395.



Protein domains bind to modified histones



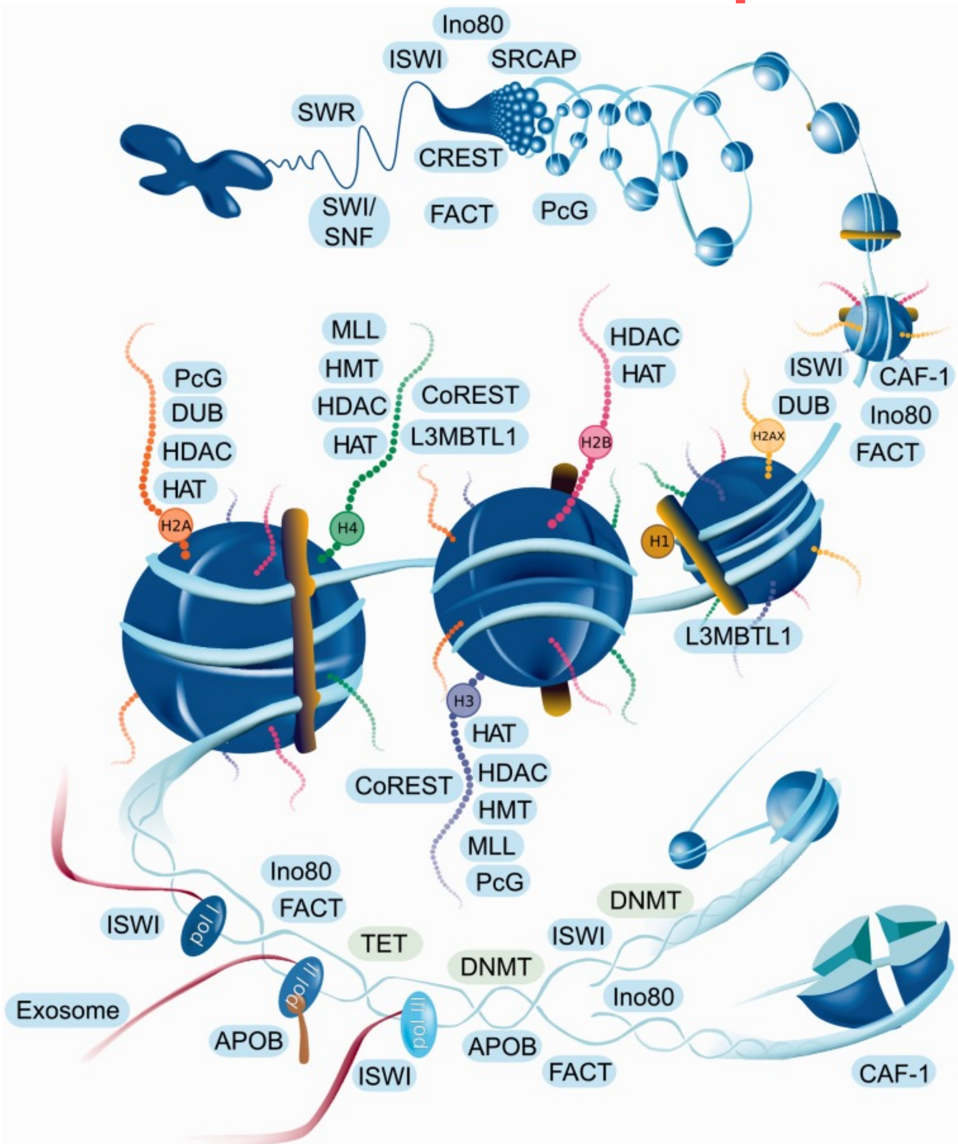
Examples of proteins with domains that specifically bind to modified histones.

There are more domain types recognizing lysine methylation than any other PTM.

H3K4me3 – a mark associated with active transcription – is recognized by a PHD finger within the ING family of proteins (ING1-5). The ING proteins in turn recruit additional chromatin modifiers such as HATs and HDACs.

Bannister, Kouzarides
Cell Res. (2011) 21: 381–395.

Epifactors database



Side view shows two windings of DNA and two histone layers

The database EpiFactors stores detailed and curated information about 815 proteins and 69 complexes involved in epigenetic regulation.

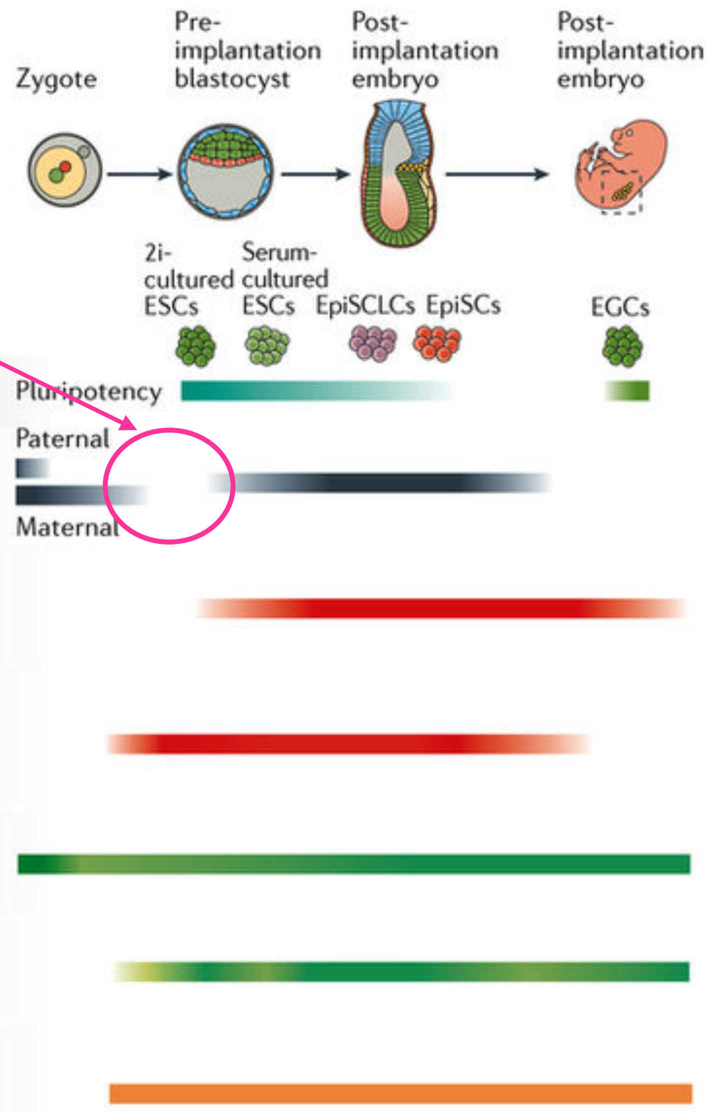
http://epifactors.autosome.ru/protein_complexes

Database (Oxford). 2015; 2015: bav067.

Dynamics of epigenetic modifications

DNA methylation is erased in the paternal and maternal genomes after fertilization and is put back on at later developmental stages.

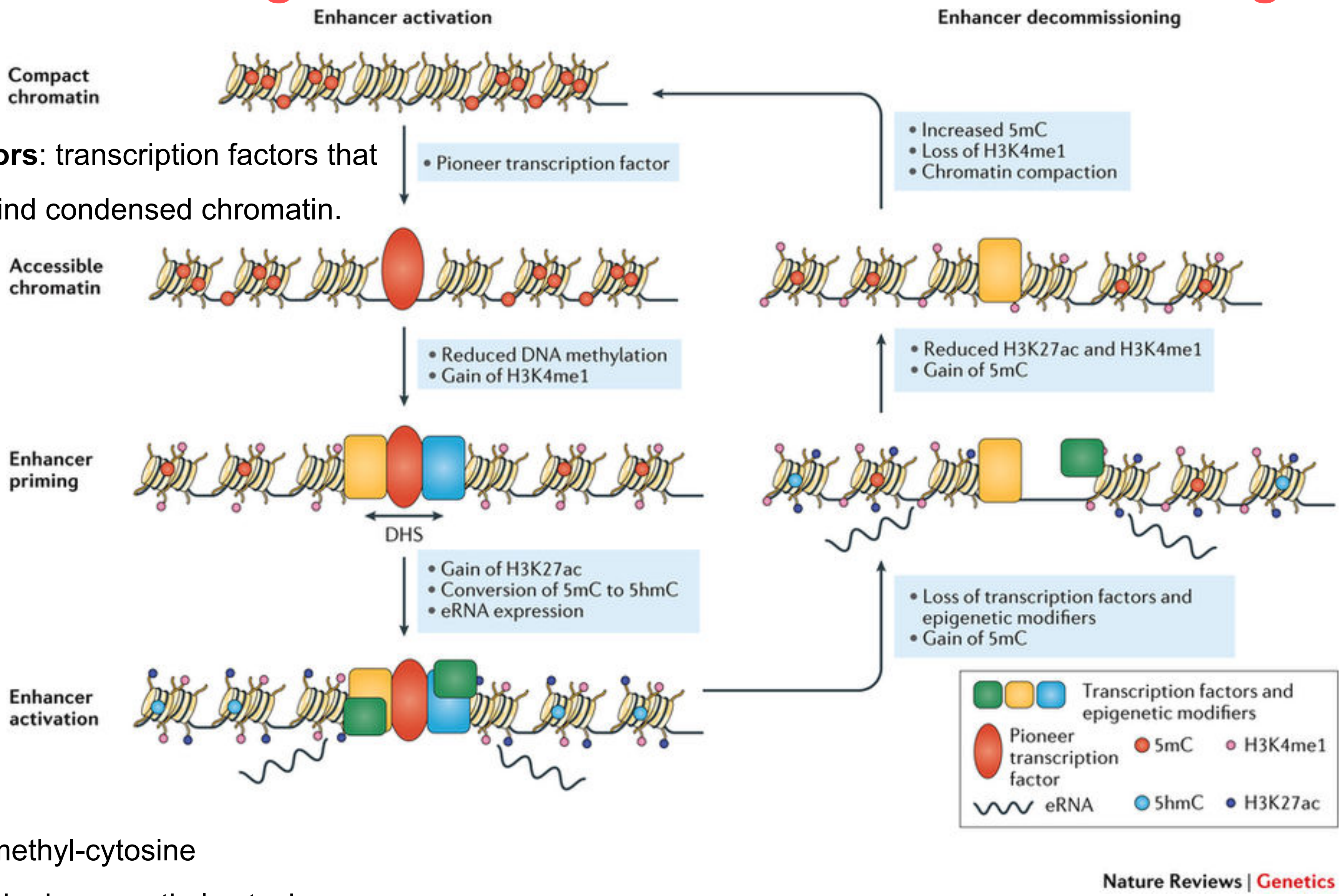
Chromatin modification	Writers	Erasers	Location	Function
DNA methylation	DNMT1, DNMT3A and DNMT3B	TET1, TET2 and TET3	CpG dinucleotides	Silencing and others
H3K27me3	PRC2	• UTX1 • JMJD3	CpG-rich promoters and intergenic regions	Silencing
H3K9me2	G9A and GLP	• JMJD2A, JMJD2B, JMJD2C and JMJD2D • JMJD1A, JMJD1B and JMJD1C	Gene bodies, intergenic regions and enhancers	Silencing
H3K4me3	COMPASS-like proteins (SET1, MLL1–MLL2)	• JARID1A, JARID1B, JARID1C and JARID1D • KDM2B	Mainly promoters	Possibly activating
H3K27ac	HATs (including CBP/p300, GNATs and MYSTs)	HDACs and sirtuins	Promoters and enhancers	Activating
H3K4me1	COMPASS-like proteins (MLL3–MLL4)	LSD1 and LSD2	Promoters, enhancers and intergenic regions	Priming and/or activating



Atlasi & Stunnenberg, *Nature Rev Genet* **18**, 643–658 (2017)

Events during enhancer activation / decommissioning

Pioneer factors: transcription factors that can directly bind condensed chromatin.

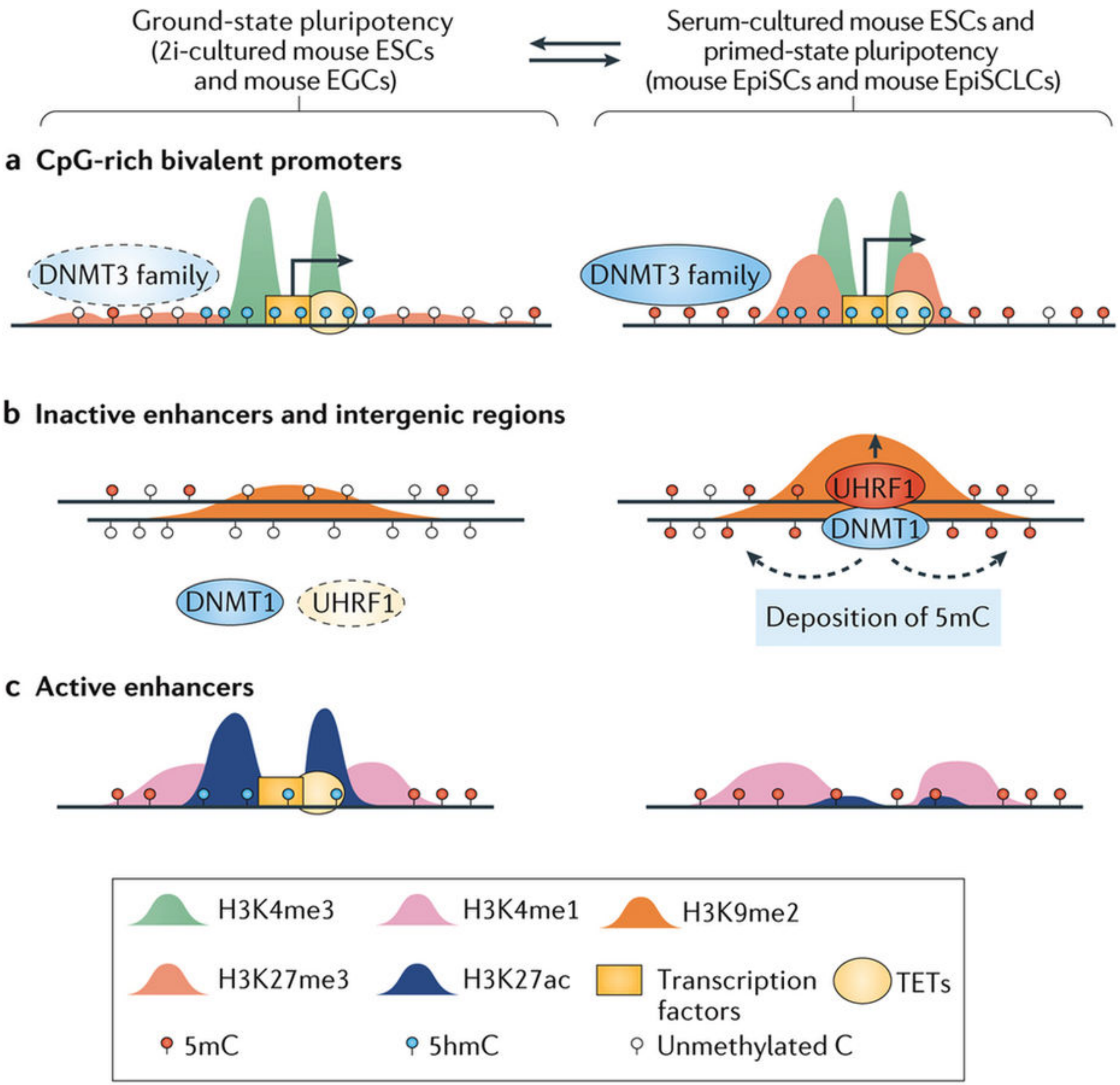


5mC: 5-methyl-cytosine

5hmC: 5-hydroxy-methyl-cytosine

Atlasi & Stunnenberg, *Nature Rev Genet* **18**, 643–658 (2017)

Interplay between DNA methylation and histone modifications



Bivalent chromatin are segments of DNA, bound to histone proteins, that have both repressing and activating epigenetic regulators in the same region. These regulators work to enhance or silence the expression of genes. Since these regulators work in opposition to each other, they normally interact with chromatin at different times. However, in bivalent chromatin, both types of regulators are interacting with the same domain at the same time. Bivalent chromatin domains are normally associated with promoters of transcription factor genes that are expressed at low levels. Bivalent domains have also been found to play a role in developmental regulation in pluripotent embryonic stems cells, as well as gene imprinting.

Atlasi & Stunnenberg, *Nature Rev Genet* **18**, 643–658 (2017)
www.wikipedia.org

Paper #7

http://www.pnas.org/content/111/26/9503.full

Molecular ties between the cell cycle and differentiation in embryonic stem cells
Victor C. Li and Marc W. Kirschner
PNAS 111, 9503–9508 (2014)

Paper presentation Jan. 8, 2018